Data Systems and Preservation Working Group Report

Paul Wagner

September 1, 2016

Aura Science Meeting – DSWG/DSPWG

Rotterdam, NL

Jet Propulsion Laboratory

California Institute of Technology

Topics

- ESDIS Report (Alfreda Hall)

- Instrument Team Ground Data System Reports
    - TES (Scott Gluck)
    - MLS (Paul Wagner, Amy Chen, Elmain Martinez, Brian Knsop)
    - OMI (Jacques Claas, Phil Durbin)

- Aura HDF-EOS Guidelines

- GES DISC Status and Data Preservation Efforts (James Johnson)

- Evolving Metadata Requirements (Ted Habermann)
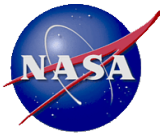
ESDIS report

Policies on openness and cooperation

- Data are openly available to all and free of charge
- NASA-developed software can be open-sourced quickly enough to allow easy collaboration and reuse
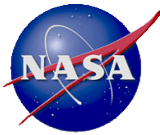
Data management services

- SIPS (Production)
- Worldview and GIBS (Discovery and visualization)
- LANCE (near-real time)

Working Groups

- Dataset Interoperability Working Group
- Data Quality Working Group

Other technology directions

- Cloud computing environments (prototyping and feasibility studies)
- Enhancing the usefulness of metadata

TES

- Failure of Laser B requires adopting SIMCLK technique
- Standard processing carried out by SIPS
  - Release 14.o (Level 2 ESDT Version 7)
  - Levels 1B, 2, 2 Lite, 3, Joint TES-MLS, and TES-OMI
- Under development (besides SIMCLK)
  - Joint TES-MLS CO and ozone based on MLS v4
  - Joint AIRS-OMI ozone
  - New ethylene, acetone, and HCN standard products
- Processing configurations
  - SIPS: 204 nodes in total, all dual Xeon quad cores
  - SCF: 50 dual Xeon quad-cores
  - TESROCKS: 124 nodes in total, all dual Xeon quad-cores
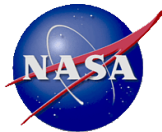
MLS

- Products and current versions
  - Standard level 2, v4.22 (replacing v3.43 and v2.33)
  - Improved CO, ozone, and OH; added DOIs and new metadata
  - Level 3 BrO and HO2; Near-real time still at v3.40
- Under development
  - Near-real time based on v4 code base, adding DOIs and new metadata
  - Possible treatment of H2O trends
- SIPS
  - Processed v3 from Aug 2009 through July 2015
  - Level 3 products beginning July 2016
  - NASA studying feasibility of terminating lease on computing facility: may insist on move to pod-based computing
- First delivery of preserved documents complete; second planned
- Services provided MLS data users
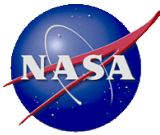
OMI

- Standard products still at Data Collection 3
- Row anomaly flagged at Level 1
  - Affected ground pixels identified using flags
  - In changes to the identified pixels, the dates will be post-processed in level 1; last update was Autumn 2014
- Analysis of all Level 1B data underway searching for trends, errors, or omissions in calibration parameters
- So far no trend suggesting need for a new Collection 4
- Since last meeting, improvements to slant height for NO2, new algorithm for SO2, a new joint OMI-MODIS Cloud product
- Plan to convert science applications to 64-bit as they are delivered
- Plan to switch Fortran compiler from PGI to gfortran
- Preservation efforts yet to begin

Aura HDF-EOS Guidelines

- Hope to create a single link to the latest version, possibly on the NASA Aura website
- Guidelines should capture changes noted in each DSWG/DSPWG breakout session; a living document
- Next version will include
  - Evolving Metadata as suggested by experience with ISO
  - New standard and joint products
  - Improved attribute naming to agree with standard practice of other satellites
  - Aim not just to help users discover data but to understand it

GES DISC Status and Data Preservation

- Implemented new Earthdata login in July 2016
  - Requires registering with a username and password as well as authorization to access data
  - This allows EOSDIS to better track data usage and understand users' needs
  - Data are still free of charge!
  - Anonymous ftp access will end October 3, 2016; users must use http access afterwards (can automate with wget/curl scripts)
- Migrated all metadata from ECHO and GCMD to new Common Metadata Repository (CMR)
- Overhauling website to a simplified Unified User Interface (UUI) combining data products, documentation, and services
- Awaiting additional MLS data preservation docs, and start of OMI data preservation
- More details to presented in Jennifer Wei's talk and James Johnson's poster

Evolving Metadata Requirements

US GEO Framework

- ISO Conceptual Model 19115-1, XML Schema 191153, and Data Quality 19157 should replace 19115-2
- We should help users not merely discover data, but to use and understand it
- Include ISO concepts and contents usefully and consistently across products and missions
- Many single ECHO Parameter name=value pairs will be replaced with groups to allow more information, sharing of content, and reuse of components across products

Tools for designers and producers

- Earthdata wiki
- Product Designer